Contents

1 Review of Computational Intelligence for Gene-Gene and Gene-Environment Interactions in Disease Mapping
Arpad Kelemen, Yulan Liang and Athanasios Vasilakos .......................... 1
1.1 Introduction .................................................. 1
1.2 Computational challenges in genomic association study of complex diseases .............................................. 3
1.3 The promise of computational intelligence ................................. 3
1.4 Computational intelligence approaches for SNP-disease associations ... 7
1.5 Conclusions .................................................................. 11
References ........................................................................... 11

2 Intelligent Approaches to Mining the Primary Research Literature:
Techniques, Systems, and Examples
Gully A.P.C. Burns, Donghai Feng, Eduard Hovy ............................... 17
2.1 Introduction .................................................. 17
2.2 A framework for conceptual biology ...................................... 19
2.3 Partitioning the literature - the notion of ‘experimental type’ ....... 23
2.4 Related Work: Biomedical Knowledge Bases based on published studies ......................................................... 25
2.5 Practical applications: ‘Research Synthesis’ ......................... 28
2.6 An Example Experiment Type: ‘Tract-Tracing Experiments’ ....... 30
2.7 Methodology ................................................................ 33
The overall challenge, from text to knowledge ................................ 33
Information extraction techniques: from Patterns to Conditional Random Fields ......................................................... 34
Stage 1: Corpus Processing ................................................. 37
Stage 2: The basic text processing and feature definition .............. 38
Stage 3: Evaluating the results ........................................... 39
2.8 Results ....................................................................... 40
2.9 Conclusions .................................................................. 42
2.10 Acknowledgements ...................................................... 43
References ........................................................................... 43
3 Intuitionistic Fuzzy Set: Application to Medical Image Segmentation
Tamalika Chaira, Tridib Chaira ........................................ 51
3.1 Introduction .................................................. 52
3.2 Related work in medical informatics ......................... 54
3.3 Preliminaries on intuitionistic fuzzy set ...................... 56
3.4 Use of intuitionistic fuzzy set theory in medical informatics -
  segmentation .................................................. 57
  3.4.1 Edge detection ......................................... 57
  3.4.2 Calculation of membership, non-membership degree
  and intuitionistic fuzzy index ............................. 60
  3.4.3 Application of intuitionistic fuzzy set theory in image
  thresholding ........................................... 61
3.5 The proposed method ....................................... 63
3.6 Results and Discussion ......................................... 65
3.7 Conclusions .................................................. 67
References ...................................................... 67

4 Decomposable Aggregability in Population Genetics and Evolutionary
Computations: Algorithms and Computational Complexity
Vladik Kreinovich, Max Shpak .................................... 69
4.1 What is Aggregability ....................................... 69
  Comment. ................................................ 71
4.2 Exact Aggregability: A Simple Example ..................... 71
4.3 Detecting Aggregability is Important ........................ 73
4.4 What We Do in This Chapter ................................ 74
  Comment. ................................................ 75
4.5 First Theoretical Result: Detecting Decomposable Aggregability
  Is NP-Hard .................................................. 76
  Comment. ................................................ 76
  Comment. ................................................ 76
  Comment. ................................................ 77
  Comments. ........................................... 77
4.6 Approximate Aggregability: Possible Definitions .............. 78
  Comment. ................................................ 80
4.7 Second Theoretical Result: Detecting Approximate Decomposable
  Aggregability Is Also NP-Hard ................................ 80
  Comment. ................................................ 81
4.8 Once We Find the Partition, Finding the Combinations is Feasible . 81
  4.8.1 Algorithm ............................................ 81
  4.8.2 Calculating aggregate dynamics ........................ 82
  4.8.3 Example .............................................. 82
  4.8.4 What to Do in the Case of Approximate Aggregation ...... 82
4.9 Proofs ....................................................... 83
7 Auxiliary tool for the identification of genetic coding sequences in eukaryotic organisms

Vincenzo De Roberto Junior Nelson F. F. Ebecken, Elias Restum Antonio . . . 133
7.1 Introduction .................................................. 133
7.2 The main gene finding programs ................................ 134
7.3 Data characteristics ........................................... 137
7.3.1 Data volume ............................................. 137
7.3.2 Information consistency .................................. 138
7.3.3 Stored information ....................................... 138
7.3.4 Database format ....................................... 138
7.4 Accuracy measures ........................................... 139
7.4.1 Nucleotide level ....................................... 139
7.4.2 Exon level ............................................ 139
7.5 Gene model .................................................. 141
7.5.1 Data selection ......................................... 142
7.5.2 The model for the discovery of coding areas .............. 142
7.6 ExonBR system ................................................ 144
7.7 Results and conclusion ....................................... 147
References ...................................................... 148

8 Language engineering and information theoretic methods in protein sequence similarity studies

A. Bogan-Marta, A. Hategan, I. Pitas .................................. 151
8.1 Introduction .................................................. 151
8.2 A survey of language engineering and information theoretic methods for protein sequence similarity ................................. 154
8.2.1 Protein sequence similarity based on linguistic techniques . . . 155
  Similarity due to regions as index terms in information retrieval 155
  Semantic similarity measures .................................. 156
  Language techniques for protein sequence relations . . . 158
8.2.2 Sequence similarity based on information theoretic methods . 160
8.3 Statistical language modeling method for sequence similarity . . 163
8.3.1 Theoretical concepts .................................. 163
8.3.2 Method description .................................. 165
8.3.3 Protein similarity search ................................ 165
  .................................................................. 166
8.3.4 Experimental results .................................. 166
8.3.5 Discussions ........................................... 168
8.4 Protein similarity detection based on lossless compression of protein sequences .................................................. 168
8.4.1 Theoretical concepts .................................. 169
8.4.2 ProtComp algorithm and its use in sequence comparison .. 169
8.4.3 Experimental results .................................. 173
8.4.4 Discussions ........................................... 175
9 Gene Expression Imputation Techniques for Robust Post Genomic Knowledge Discovery

Muhammad Shoaib Sehgal, Iqbal Gondal and Laurence Dooley

9.1 Introduction .................................................. 185
9.2 Missing Values Estimation Methods .......................... 187
9.2.1 Imputation Nomenclature .................................. 187
9.2.2 Zero and Average Based Imputation [17] ................. 189
9.2.3 Singular Value Decomposition Based Imputation (SVDImpute) [12] ........................................... 189
9.2.4 K-Nearest Neighbour (KNN) Estimation [12] ............. 190
9.2.5 Least Square Impute (LSImpute) [13] ................. 191
9.2.6 Local Least Square Impute (LLSImpute) [10] .......... 191
9.2.7 Bayesian Principal Component Analysis based Estimation (BPCA) [6] ........................................... 192
9.2.8 Collateral Missing Value Estimation (CMVE) [14] .... 192
9.3 Post Genomic Knowledge Discovery Methods .............. 194
9.3.1 Gene Selection ........................................ 195
9.3.2 Gene Regulatory Network Reconstruction ................. 195
9.3.3 Statistical Significance Test ............................. 196
9.4 Analysis and Discussion of Results .......................... 196
9.5 Software Availability ........................................ 203
9.6 Conclusions .................................................. 203
References ...................................................... 203

10 Computational Modelling Strategies for Gene Regulatory Network Reconstruction

Muhammad Shoaib Sehgal, Iqbal Gondal and Laurence Dooley

10.1 Introduction .................................................. 207
10.2 Pair Wise GRN Reconstruction Methods ..................... 210
10.2.1 Correlation and Distance Functions ..................... 211
10.2.2 Mutual Information .................................... 211
10.2.3 Local Shape Based Similarity .......................... 212
10.3 Deterministic Methods for GRN Inference ..................... 213
10.3.1 Differential Equations ................................. 213
10.3.2 Boolean GRN Modelling Methods ....................... 213
10.4 Probabilistic GRN Reconstructed Strategies ................. 214
10.5 Hybrid GRN Inference Methods ............................. 215
10.6 Conclusions .................................................. 217
References ...................................................... 218
11 Integration of Brain-Gene Ontology and Simulation Systems for Learning, Modelling and Discovery
Nik Kasabov, Vishal Jain, Lubica Benuska, Paulo C. M. Gottgtroy and Frances Joseph ................................................ 221
11.1 Introduction .................................................. 221
11.2 Evolving Implementation of BGO in Protégé ....................... 223
11.3 New Discoveries with BGO ..................................... 225
11.4 BGO for Teaching ............................................ 231
11.5 Conclusions and Future Directions ................................ 231
11.6 Acknowledgments ............................................. 233
References .................................................................. 233

12 Efficiency and Scalability Issues in Metric Access Methods
Vlastislav Dohnal, Claudio Gennaro, Pavel Zezula .................... 235
12.1 The Metric Space Approach ...................................... 236
12.1.1 Metric Spaces in Bioinformatics ............................. 236
12.1.2 Software Tools ........................................ 238
12.2 Centralized Metric Access Methods ................................ 239
12.2.1 Metric Tree Family ..................................... 239
12.2.2 Hash-based Similarity Indexing ............................ 241
12.2.3 Performance Trials ........................................ 241
12.3 Exploiting Multiple Computational Resources ..................... 243
12.3.1 Parallel Similarity Searching ................................ 244
12.3.2 Centralized Coordination .................................... 244
  Metric Grid .................................................. 244
12.3.3 Scalable and Distributed Search Structures ............... 245
  Native Metric Search Structures ................................. 246
  Metric Content Addressable Network ........................... 248
  Metric Chord ........................................... 250
12.4 Experience from Performance Trials .............................. 251
12.4.1 Performance Tuning ...................................... 254
12.5 Conclusions .................................................. 257
References .................................................................. 259

13 Computational Modelling of the Biomechanics of Epithelial and Mesenchymal Cell Interactions During Morphological Development
Jiří Kroc .......................................................... 265
13.1 Introduction .................................................. 265
13.2 Introduction to Cell and Tissue Biology ............................ 268
13.3 Overview of Morphological Development of Tooth Resulting from Interaction of Epithelium and Mesenchyme ................. 269
13.4 Simplifications used in model of morphological growth ........... 271
13.5 Model .......................................................... 273
  13.5.1 Three inter-dependent parts of the model ................. 273
  13.5.2 Mechanical interaction of epithelium and mesenchyme 274
Contents 

13.5.3 Tissue growth .............................................. 278
13.5.4 Implementation of diffusion by two computational sub-steps . 279
13.5.5 Signalling, switching and cell differentiation .................. 280
13.6 Results and Discussions ...................................... 281
13.7 Future development .......................................... 291
13.8 Conclusions .................................................. 292
References .......................................................... 293

14 Artificial Chemistry and Molecular Darwinian Evolution 
of DNA/RNA-Like Systems I - Typogenetics and Chemostat 
Marian Bobrik, Vladimir Kvasnicka, Jiri Pospichal ................... 295

14.1 Introduction .................................................. 295
14.2 Eigen Theory of Molecular Replicators ........................... 298
14.2.1 Replicators and Molecular Darwinian Evolution ................ 301
14.2.2 Artificial Chemistry and a Metaphor of Chemostat .......... 303
14.2.3 Summary ................................................... 306
14.3 Hofstadter’s Theory of DNA-Like Molecules, Typogenetics .... 307
14.3.1 Basic Principles of Typogenetics ................................ 309
14.3.2 An Expression of Strands by Enzymes ....................... 310
14.3.3 Replicators ................................................ 314
14.3.4 Hypercycles .............................................. 319
14.3.5 Summary ................................................... 321
14.4 Folding of RNA-Like Molecules that are Represented by Binary Strings 321
14.4.1 Chemostat Simulation of Molecular Darwinian Evolution .... 330
14.4.2 Summary ................................................... 332
References .......................................................... 334

15 Artificial Chemistry and Molecular Darwinian Evolution 
of DNA/RNA-Like Systems II – Programmable folding 
Marian Bobrik, Vladimir Kvasnicka, Jiri Pospichal ................... 337

15.1 Programmable Folding of RNA-Like Molecules .................. 337
15.2 Source of the Instruction Set .................................. 338
15.3 Architecture of Programmable RNA System ...................... 342
15.4 Comparison with the Instruction Based System ................. 347
15.5 Replicator Evolution by Artificial Selection ..................... 348
15.5.1 The behavior of the System ................................ 350
15.5.2 String Splitting .......................................... 354
15.6 The Chemostat ................................................. 358
15.6.1 Perfectly Mixed Chemostat .................................. 360
15.6.2 Chemostat with Spatial Structure ........................... 360
15.6.3 Replicator Spatial Structures ................................ 362
15.6.4 The Mechanism of Resistance against Parasites .......... 366
15.7 Conclusions .................................................. 371
References .......................................................... 372

Index .............................................................. 375
Computational Intelligence in Medical Informatics
Kelemen, A.; Abraham, A.; Liang, Y. (Eds.)
2008, XV, 379 p., Hardcover
ISBN: 978-3-540-75766-5